Multiband treatment of quantum transport in interband tunnel devices

D. Z.-Y. Ting, E. T. Yu, and T. C. McGill
Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125
(Received 13 June 1991)

We describe a method for computing transmission coefficients for multiband tight-binding band-structure models. In this method, the transmission probability can be calculated simply by solving a system of linear equations representing the tight-binding form of the Schrödinger equation over a finite region of interest, with specially formulated boundary and inhomogeneous terms to account for the effects of the incoming and outgoing plane-wave states. In addition to being efficient, and simple to implement, our method is numerically stable in treating device structures with large active regions, and therefore capable of modeling realistic band-bending effects. Using this method, we examine transport properties in InAs/GaSb/AlSb-based interband tunnel structures with a realistic band-structure model. We compare our results with calculations obtained with a two-band model, which includes only the lowest conduction band and the light-hole band. We find that while the primary interband transport mechanism arises from the coupling between the InAs conduction-band states and GaSb light-hole states, in device structures containing GaSb quantum wells, the inclusion of heavy-hole states can introduce additional transmission resonances and substantial hole-mixing effects. These effects are found to have a significant influence on the current-voltage characteristics of interband devices.

I. INTRODUCTION

The nearly lattice-matched InAs/GaSb/AlSb material system has received considerable attention recently due to the tremendous flexibility it offers for heterostructure device design. As shown in Fig. 1, in which the conduction- and valence-band edges for InAs, GaSb, and AlSb are plotted according to the currently accepted band offset values,[1-3] this material system offers the possibility for type-I (GaSb-AlSb), type-II staggered (AlSb-InAs), and type-II broken-gap (InAs-GaSb) band alignments. Recently, several interband devices, so called because they exploit the type-II broken-gap band alignment between InAs and GaSb, have been demonstrated.[4-11] Various interband devices have exhibited very high-peak current densities or large peak-to-valley current ratios, making them extremely attractive for use in high-frequency oscillators, logic circuits, and a variety of other digital and analog applications. For example, a peak-to-valley current ratio of 20 (88) at 300 K (77 K) has been reported in an InAs-AlSb-GaSb-AlSb-InAs device,[4] and negative differential resistance (NDR) with peak current densities in excess of $10^5 A/cm^2$ has been observed in a number of different interband devices.[10] In addition, GaSb-InAs-AlSb-GaSb Stark effect tunneling transistors utilizing interband transport properties have demonstrated high gains at room temperature.[11]

To date, most of the theoretical studies of interband tunneling have used simple two-band models which included only the interaction between the conduction and light-hole bands, and assumed that interactions involving the heavy-hole and split-off bands are negligible.[12,13] Although this simple approach has been fairly successful in explaining qualitative features observed experimentally,[14] it is somewhat unsatisfactory due to the oversimplification of the valence-band structure. Recently, we reported briefly on a more realistic calculation in which we showed that the calculated transmission coefficients in interband tunnel structures can be significantly altered when heavy-hole states are taken into account.[15] In this paper we give a detailed account

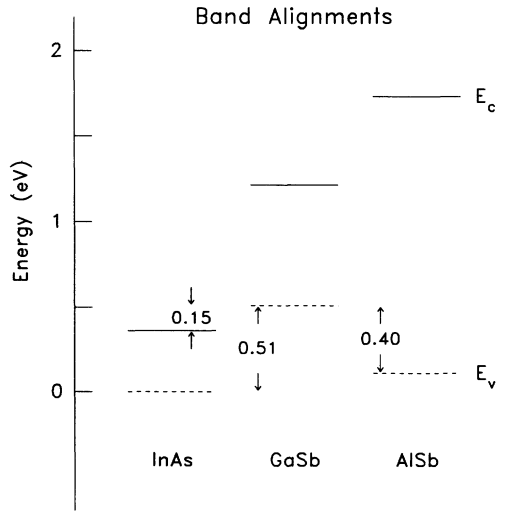

FIG. 1. Conduction- (solid) and valence- (dashed) band edges for the InAs/GaSb/AlSb material system. The energy gaps and band offsets allow the possibility of type I, type II, and type II broken-gap band alignments. The indirect conduction-band minimum is shown for AlSb.

of the methods used in our calculation, and report on the calculation of transmission coefficients and current-voltage characteristics for several prototypical interband devices.

To incorporate realistic valence-band structure in our treatment of interband tunnel devices, we have used the eight-band effective bond-orbital model[16] developed by Chang. This model is essentially a reformulation of Kane's eight-band $\mathbf{k}\cdot\mathbf{p}$ model[17] in the tight-binding framework. For tunneling calculations, the effective bond-orbital model offers two advantages over the $\mathbf{k}\cdot\mathbf{p}$ model: (1) It is more efficient numerically since the the matrix elements relating values of the wave function in adjacent regions (hopping matrix elements) contain simple algebraic expressions rather than transcendental functions, and (2) it allows for a much simpler treatment of boundary conditions at the heterointerfaces.

A problem that arises in transport calculations incorporating realistic band structures is the need for a numerically stable algorithm for calculating transmission coefficients. A standard method for computing transmission coefficients in heterostructures is the transfer-matrix method.[18,19] It is well known, however, that the transfer-matrix method, when used in conjunction with realistic multiband band-structure models, is numerically unstable for treating device structures with active regions larger than a few tens of Å.[20] When band-bending effects are included in the calculations of current-voltage characteristics, it is typically necessary to compute transmission coefficients for structures considerably wider than 1000 Å. There are several numerically stable methods for calculating transmission coefficients.[21–25] Among them is a method by Frensley[23] which uses an approach similar to that of Lent and Kirkner in their treatment of quantum waveguides.[26] Frensley originally used his method with the single-band envelope-function model to calculate transmission coefficients in tunnel structures. We have generalized this method to multiband tight-binding models to treat band-structure effects in a wide variety of heterostructure tunnel devices. We have implemented this method for the effective bond orbital model[16] in our treatment of interband tunneling in InAs/GaSb/AlSb heterostructures,[15] and hole tunneling[27] in GaAs/AlAs double barrier heterostructures. We have also implemented this method for the second-neighbor sp^3 tight-binding model[28] in treating X-point tunneling[29] in GaAs/AlAs double barrier heterostructures. The method has been demonstrated to be numerically stable in our transmission coefficient calculations for device structures wider than 2000 Å. In addition, our analysis indicates that the computational cost of this method is linearly proportional to the width of the heterostructure. In fact, this method is as efficient as the transfer-matrix method. Furthermore, our method also has the advantage of being simpler to implement than the transfer-matrix method.

In Sec. II of this paper, we describe the methods we used to model interband tunnel devices. In Sec. III, we present our analysis of several prototypical InAs/GaSb/AlSb interband tunnel devices. A summary is given in Sec. IV.

II. METHODS

Section II A outlines the procedure used to model interband tunnel devices, including the treatment of band-bending effects, the choice of band-structure model, and the calculation of current-voltage characteristics. We also present a detailed discussion of the method that we have developed to compute transmission coefficients for multiband tight-binding band-structure models. Although the general formulation of the transmission coefficient calculation resembles the transfer-matrix method to a certain extent,[19] the solution of the problem is very different. Section II B discusses the formulation of the transmission coefficient calculation, and Sec. II C presents a description and analysis of our solution.

A. General procedure

Our calculation of current-voltage characteristics consists of four major components. First, band bending corresponding to the given doping profile and applied bias is computed to give the energy-band profile. At the same time, a suitable band-structure model is chosen to realistically describe the materials comprising the heterostructure. Next, transmission coefficients are computed for device profiles obtained from the band-bending calculation. Finally, current density is computed from the transmission coefficients. Our calculation does not take into account inelastic-scattering processes, nor does it require self-consistency between the band profile and the wave functions of the transmitting states. The emphasis of this calculation is to examine the role of realistic band structures in describing the operations in interband tunnel devices.

The first step in our simulations of heterostructure devices is to calculate the energy-band profile of the structure. We compute the energy-band profile by solving Poisson's equation across the device, imposing a condition of charge neutrality over the entire device structure. The Thomas-Fermi approximation is used to relate the positions of the conduction- and valence-band edges at a given point to the local carrier concentration, and the effects of finite temperature on the Fermi distribution are included. For the band-bending calculations, a simple parabolic band-structure model is typically assumed for the conduction and valence bands. The actual numerical calculation of the heterostructure band-edge profile is performed by converting Poisson's equation to a finite-difference equation and applying a relaxation algorithm described by Press *et al.*[30] The energy-band profile of a typical InAs-GaSb-InAs device obtained from the band-bending calculation is shown in Fig. 2. In this example, the width of the GaSb quantum well is 40 monolayers, and the doping level in the InAs electrodes is $n = 10^{17}$ cm^{-3}. The device is under a 100 mV applied bias.

To include band-structure effects properly, we use the eight-band effective bond-orbital model,[16] which incorporates spin-orbit interaction and accurately describes the heavy-hole, light-hole, and split-off valence bands and the lowest conduction band near the center of the Brillouin zone. The basis set contains eight effective bond

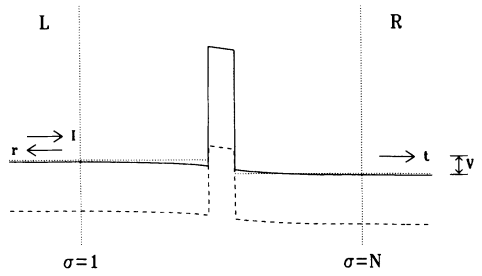

FIG. 2. Energy-band diagram of an InAs-GaSb-InAs interband tunnel structure under an apllied bias of 100 mV. The width of the GaSb quantum well is 40 monolayers, and doping level in the InAs electrodes is n=10^{17} cm^{-3}.

orbitals per unit cell. Six of the orbitals are obtained by taking linear combinations of p-like orbitals coupled with electron spin; they form states with total angular momentum of $J = \frac{3}{2}$ ($J_z = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$), or $J = \frac{1}{2}$ ($J_z = -\frac{1}{2}, \frac{1}{2}$). The remaining two orbitals are s-like orbitals with spin up and spin down. At the Brillouin-zone center, the $|\frac{3}{2}, \pm\frac{3}{2}\rangle$, $|\frac{3}{2}, \pm\frac{1}{2}\rangle$, $|\frac{1}{2}, \pm\frac{1}{2}\rangle$, and $|s, \pm\frac{1}{2}\rangle$ orbitals correspond to the heavy-hole, light-hole, and split-off valence bands and the lowest conduction band, respectively. The effective bond-orbital model provides a realistic description of the relevant band structure, and takes advantage of the flexibility and efficiency of the tight-binding method.

Having chosen a suitable band-structure model to work with, transmission coefficients are computed for device profiles obtained from the band-bending calculation. The general prescription for calculating the current density from the transmission coefficient has been adapted from the approach of Duke.[31] Briefly stated, the method consists of integrating over the Fermi distribution of the incident electron population, including appropriate Fermi factors for occupied states in the emitter and empty states in the collector, and including appropriate velocity factors for the incident electrons. The total energy E and the wave-vector component parallel to the device $\mathbf{k}_\parallel$ are taken to be conserved. The current density $\mathcal{J}$ is then given by[31]

$$\mathcal{J} = \frac{e}{4\pi^3\hbar} \int T(E, \mathbf{k}_\parallel)[f(E) - f(E + eV)] dE_\perp d^2k_\parallel, \tag{1}$$

where V is the bias voltage applied to the device structure, and $T(E, \mathbf{k}_\parallel)$ is the transmission coefficient. A detailed description of the method we used to compute transmission coefficients is given below.

B. Formulation of transmission coefficient calculation

We divide the heterostructure into three sections as shown in Fig. 2: a semi-infinite flat-band region on the left, a semi-infinite flat-band region on the right, and a central N-monolayer segment containing the heterointerfaces. The left and right sections are chosen to be regions with constant doping levels that are sufficiently far away from the heterointerfaces so that the potentials and compositions are constant; the wave function in these regions can therefore be described in terms of bulklike plane-wave states. In general, the objective of transmission coefficient calculations is to relate the amplitude of the incoming plane-wave states (from the left, in our case) to the amplitudes of the transmitted plane-wave states (to the right). In the tight-binding framework, the strategy is to translate the amplitudes of the plane-wave states into coefficients of tight-binding orbitals, and relate the coefficients in the right and left regions through tight-binding calculations. The additional basis transformation yields much simpler hopping matrix elements, and allows the computation to be performed much more efficiently.

We consider the heterostructure as a sequence of monolayers parallel to the heterointerfaces. Let M be the number of orbitals per unit cell in our tight-binding basis set ($M = 8$ in our case). The basis orbitals may be written in the form $|\mathbf{R}_\parallel \sigma\alpha\rangle$, where σ is an integer monolayer label ($\sigma = 1, 2, ..., N$ in the central region), $\mathbf{R}_\parallel$ specifies the in-plane component of unit cell coordinate, and $\alpha = 1, 2, ..., M$ labels the orbitals within a unit cell. Since the in-plane crystal momentum, $\mathbf{k}_\parallel$, is a good quantum number, the wave function may be written as

$$|\psi\rangle = \sum_{\sigma,\alpha} C_{\sigma\alpha} |\sigma\alpha, \mathbf{k}_\parallel\rangle, \tag{2}$$

where $|\sigma\alpha, \mathbf{k}_\parallel\rangle$ is a planar orbital formed by taking Bloch sums of tight-binding orbitals over the $N_\parallel$ unit cells in the σth monolayer :

$$|\sigma\alpha, \mathbf{k}_\parallel\rangle = \frac{1}{\sqrt{N_\parallel}} \sum_{\mathbf{R}_\parallel} \exp(i\mathbf{k}_\parallel \cdot \mathbf{R}_\parallel) |\mathbf{R}_\parallel \sigma\alpha\rangle. \tag{3}$$

Writing the Schrödinger equation $(H - E)|\psi\rangle = 0$ in the planar orbital basis we obtain

$$\mathbf{H}_{\sigma,\sigma-1}\mathbf{C}_{\sigma-1} + \bar{\mathbf{H}}_{\sigma,\sigma}\mathbf{C}_\sigma + \mathbf{H}_{\sigma,\sigma+1}\mathbf{C}_{\sigma+1} = \mathbf{0}, \tag{4}$$

where $\mathbf{C}_\sigma$ is a vector of length M,

$$\mathbf{C}_\sigma = \begin{pmatrix} C_{\sigma 1} \\ C_{\sigma 2} \\ \vdots \\ C_{\sigma M} \end{pmatrix} \tag{5}$$

and $\mathbf{H}_{\sigma,\sigma'}$ and $\bar{\mathbf{H}}_{\sigma,\sigma}$ are $M \times M$ matrices whose elements are given by, respectively,

$$(\mathbf{H}_{\sigma,\sigma'})_{\alpha,\alpha'} = \langle \sigma\alpha, \mathbf{k}_\parallel | H | \sigma'\alpha', \mathbf{k}_\parallel \rangle \tag{6}$$

and

$$(\bar{\mathbf{H}}_{\sigma,\sigma})_{\alpha,\alpha'} = \langle \sigma\alpha, \mathbf{k}_\parallel | (H - E) | \sigma\alpha', \mathbf{k}_\parallel \rangle. \tag{7}$$

For convenience, we have suppressed the dependence on $\mathbf{k}_\parallel$ and E in our notation.

We have assumed that for the band structure model used, the interaction between two planar orbitals does not extend beyond the nearest layer, i.e.,

$$\mathbf{H}_{\sigma,\sigma'} = \mathbf{0} \quad \text{for} \quad |\sigma - \sigma'| > 1. \tag{8}$$

The method can be modified appropriately to handle more extended interactions.

It is common to write Eq. (4) in the transfer-matrix form,

$$\begin{pmatrix} -\mathbf{H}_{\sigma,\sigma-1}^{-1}\bar{\mathbf{H}}_{\sigma,\sigma} & -\mathbf{H}_{\sigma,\sigma-1}^{-1}\mathbf{H}_{\sigma,\sigma+1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{C}_\sigma \\ \mathbf{C}_{\sigma+1} \end{pmatrix} \equiv \mathbf{T}_\sigma \begin{pmatrix} \mathbf{C}_\sigma \\ \mathbf{C}_{\sigma+1} \end{pmatrix} = \begin{pmatrix} \mathbf{C}_{\sigma-1} \\ \mathbf{C}_\sigma \end{pmatrix}, \quad (9)$$

where $\mathbf{T}_\sigma$ is the single-layer transfer matrix.

The boundary conditions are such that we have a known incoming plane-wave state from the left region, no incoming states from the right, and unknown outgoing transmitted and reflected plane-wave states in the right and left regions, respectively. Although this is most naturally described in the plane-wave basis, it can be easily translated into the tight-binding basis following the prescription given by Schulman and Chang.[19] We shall describe this in some detail since it is relevant to the formulation of our treatment. First, the available plane wave states in the left and right regions can be found by noting that for a Bloch state

$$|k_\perp\rangle = \sum_{\sigma,\alpha} B_{\sigma\alpha}|\sigma\alpha, \mathbf{k}_\parallel\rangle, \quad (10)$$

the tight-binding coefficients must obey the relation

$$\mathbf{B}_\sigma = e^{ik_\perp d}\mathbf{B}_{\sigma-1}, \quad (11)$$

where d is the distance between monolayers, and $k_\perp$ is the component of the crystal momentum along the growth direction. This condition, together with Eq. (9), constitutes the following eigenvalue problem:[32]

$$\begin{pmatrix} -\mathbf{H}_{\sigma,\sigma-1}^{-1}\bar{\mathbf{H}}_{\sigma,\sigma} & -\mathbf{H}_{\sigma,\sigma-1}^{-1}\mathbf{H}_{\sigma,\sigma+1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{B}_\sigma \\ \mathbf{B}_{\sigma+1} \end{pmatrix} = e^{-ik_\perp d} \begin{pmatrix} \mathbf{B}_\sigma \\ \mathbf{B}_{\sigma+1} \end{pmatrix}. \quad (12)$$

Solving Eq. (12) yields a set of $2M$ complex wave vectors $\{k_{\perp,j}; j = 1, 2, ..., 2M\}$, and their associated bulk complex wave vector states. We order the wave vectors such that $j = 1, 2, ..., M$ corresponds to states which propagate or decay to the right, while $j = M+1, M+2, ..., 2M$ corresponds to states which propagate or decay to the left. The boundary conditions can be described in the bulk-state basis by choosing the proper form for the wave function in the left and right regions:[19]

$$|\psi; L\rangle = \sum_{j=1}^{M} \left(I_j|k_{\perp,j}; L\rangle + r_j|k_{\perp,j+M}; L\rangle\right), \quad (13)$$

$$|\psi; R\rangle = \sum_{j=1}^{M} t_j|k_{\perp,j}; R\rangle, \quad (14)$$

where $|k_{\perp,j}; L\rangle$ and $|k_{\perp,j}; R\rangle$ are the bulk complex band states in the left and right regions, respectively. Let $\mathbf{I}$, $\mathbf{r}$, and $\mathbf{t}$ be column vectors of length M containing the coefficients $\{I_j\}$, $\{r_j\}$, and $\{t_j\}$, respectively. $\mathbf{I}$ represents the known incoming states, while $\mathbf{r}$ and $\mathbf{t}$ describe the reflected and transmitted components. By examining Eqs. (10), (13), and (14), we find that $\mathbf{I}$, $\mathbf{r}$, and $\mathbf{t}$ can be related to the layer orbital coefficients by a simple basis transformation:

$$\begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{pmatrix} = \mathbf{D}^L \begin{pmatrix} \mathbf{I} \\ \mathbf{r} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{11}^L & \mathbf{D}_{12}^L \\ \mathbf{D}_{21}^L & \mathbf{D}_{22}^L \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ \mathbf{r} \end{pmatrix}, \quad (15)$$

$$\begin{pmatrix} \mathbf{C}_{N-1} \\ \mathbf{C}_N \end{pmatrix} = \mathbf{D}^R \begin{pmatrix} \mathbf{t} \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{11}^R & \mathbf{D}_{12}^R \\ \mathbf{D}_{21}^R & \mathbf{D}_{22}^R \end{pmatrix} \begin{pmatrix} \mathbf{t} \\ \mathbf{0} \end{pmatrix}. \quad (16)$$

$\mathbf{D}^L$ and $\mathbf{D}^R$ are $2M \times 2M$ matrices whose column vectors are the eigenvectors obtained by solving Eq. (12) with $\sigma = 1$ and $\sigma = N-1$, respectively, and arranged in the same order as the corresponding eigenvalues. We subdivide each $\mathbf{D}_{ij}$ matrix into four $M \times M$ matrices for convenience.

In the transfer-matrix method, $\mathbf{r}$ and $\mathbf{t}$ are obtained by relating Eqs. (15) and (16) to each other through the repeated application of transfer matrices described by Eq. (9). Unfortunately, in using this method with realistic multiband band-structure models, one always encounters numerical instability problems for wide device structures. Our experience in using the transfer-matrix method with the effective bond-orbital model shows that numerical instabilities typically set in as we consider structures wider than a few tens of Å.

C. Algorithm for computing transmission coefficient

In this work we introduce a different approach to circumvent the numerical instabilities encountered in the transfer matrix method. The algorithm is essentially a multiband generalization of the method of Frensley.[23] The idea is to eliminate the unknowns $\mathbf{r}$ and $\mathbf{t}$ and formulate the problem as a system of linear equations with only the tight-binding coefficients as the unknowns.

Eliminating $\mathbf{r}$ and $\mathbf{t}$ from Eqs. (15) and (16), we obtain

$$\mathbf{C}_1 - \mathbf{D}_{12}^L\mathbf{D}_{22}^{L\,-1}\mathbf{C}_2 = \mathbf{D}_{11}^L\mathbf{I} - \mathbf{D}_{12}^L\mathbf{D}_{22}^{L\,-1}\mathbf{D}_{21}^L\mathbf{I}, \quad (17)$$

and

$$-\mathbf{D}_{21}^R\mathbf{D}_{11}^{R\,-1}\mathbf{C}_{N-1} + \mathbf{C}_N = \mathbf{0}. \quad (18)$$

The above equations, together with Eq. (4), constitute a system of MN linear equations which can be written in the matrix form as

$$\begin{pmatrix} \mathbf{1} & -\mathbf{D}_{12}^{L}\mathbf{D}_{22}^{L\,-1} & \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} & \mathbf{H}_{2,3} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{3,2} & \mathbf{H}_{3,3} & \mathbf{H}_{3,4} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \\ \mathbf{0} & & & \mathbf{0} & \mathbf{H}_{N-1,N-2} & \bar{\mathbf{H}}_{N-1,N-1} & \mathbf{H}_{N-1,N} \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} & -\mathbf{D}_{21}^{R}\mathbf{D}_{11}^{R\,-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \\ \vdots \\ \mathbf{C}_{N-1} \\ \mathbf{C}_N \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{D}_{11}^{L}\mathbf{I} - \mathbf{D}_{12}^{L}\mathbf{D}_{22}^{L\,-1}\mathbf{D}_{21}^{L}\mathbf{I} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} . \quad (19)$$

The matrix equation above is simply the tight-binding form of the Schrödinger equation for the central region, with specially formulated boundary and inhomogeneous terms to represent the effects due to the plane-wave states traveling to and from the semi-infinite regions on the left and the right. Since, other than the few special terms, Eq. (19) contains just the Hamiltonian matrix elements, the implementation of this algorithm is very simple.

Equation (19) can be solved readily using numerical algorithms from standard mathematical libraries; the band matrix routines from LINPACK (Ref. 33) work quite well for this purpose. The solution takes approximately $np(p+1)$ and $n(2p+1)$ multiplications[34] for Gaussian elimination and back substitution, respectively, where n is the order of the matrix, and p is the number of upper diagonals. In our case, $n = NM$ and $p = 2M-1$, yielding approximately $4M^2N$ multiplications. (Note: This is without pivoting; pivoting algorithms take $8M^2N$ operations.) A simple analysis shows that an efficient implementation of the transfer-matrix method also has a computational cost of $4M^2N$ multiplications. Therefore, the method presented here is as efficient as the transfer-matrix method, but with the added advantage of being numerically stable and simple to implement.

Having obtained the coefficients of the layer orbitals, it follows from Eq. (16) that the coefficients of the transmitted plane-wave states are given by

$$\mathbf{t} = \mathbf{D}_{21}^{R\,-1}\mathbf{C}_N . \quad (20)$$

In turn, $\mathbf{t}$ can be used to compute the transmission coefficient :

$$T(E,\mathbf{k}_{\|}) = \sum_{j=1}^{M} |\, t_j(E,\mathbf{k}_{\|})\,|^2 \, \frac{|v_j(E,\mathbf{k}_{\|};R)|}{|v_I(E,\mathbf{k}_{\|};L)|} , \quad (21)$$

where $v_I(E,\mathbf{k}_{\|};L)$, and $v_j(E,\mathbf{k}_{\|};R)$ are the group velocities of the incident and the transmitted bulk pane-wave states, respectively.

We end this section by briefly comparing our method with some of the other available methods.[21,22,24,25] The methods of Wachutka[21] and of Ko and Inkson[22] are more suited for use with $\mathbf{k}\cdot\mathbf{p}$ and pseudopotential models. We focus on the methods of Boykin, van der Wagt, and Harris,[24] and of Schulman,[25] which are tailored for local-orbital band-structure models. These two methods are closely related, with the latter in a simpler formulation. To avoid numerical instabilities associated with computing transfer matrices for wide device structures, in their approach a tunnel structure is divided into short subregions over which transfer matrices can be successfully computed. The subregion transfer matrices are then used to formulate a system of linear equations, which can be solved to obtain transmission coefficients in a numerically stable manner. In a sense, their method can be considered a hybrid between the transfer-matrix method and our method. But since the purely linear-systems approach that we use is comparable to the transfer-matrix method in efficiency, our method is at least as efficient as the hybrid approach, and considerably easier to implement.

III. RESULTS AND DISCUSSION

In general, two-terminal interband tunnel devices can be classified according to their terminal types : (1) two n-type electrodes, (2) two p-type electrodes, and (3) one n-type and one p-type electrode. The three categories are schematically illustrated in Fig. 3 by, respectively, (a) the InAs-GaSb-InAs structure, (b) the GaSb-InAs-GaSb structure, and (c) the InAs-GaSb-AlSb-GaSb structure. As shown in Fig. 3, the conduction band of InAs overlaps in energy with the valence bands of GaSb, leading to the possibility of interband transport. We choose the InAs conduction-band edge as the origin of the energy scale, placing the GaSb valence-band edge at 0.154 eV. Below we present results for each of the three device structures depicted in Fig. 3, with emphasis on transport in the broken-gap energy range between the InAs conduction-band edge and the GaSb valence-band edge (0–0.154 eV).

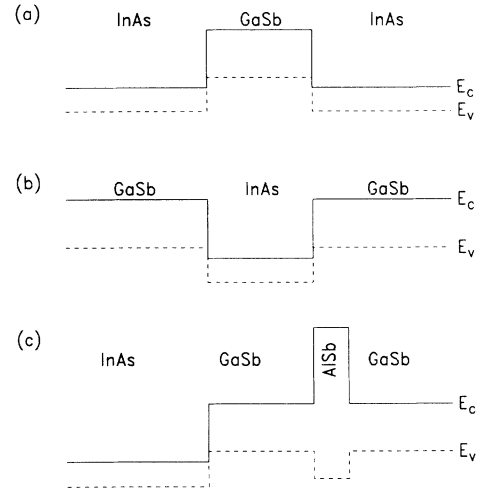

FIG. 3. Schematic energy-band diagrams for (a) the InAs-GaSb-InAs interband tunnel structure, (b) the GaSb-InAs-GaSb structure, and (c) the InAs-GaSb-AlSb-GaSb structure.

A. InAs-GaSb-InAs

Figure 3(a) shows a schematic energy-band diagram for the InAs-GaSb-InAs structure. In contrast to a double barrier heterostructure, in which resonant tunneling occurs via quasibound states localized by the barriers, quasibound hole states in the InAs-GaSb-InAs structure are formed in the GaSb layer due to the imperfect matching of InAs conduction-band and GaSb valence-band wave functions at the two InAs/GaSb interfaces. Consequently, resonant transport in this device can occur despite the absence of classically forbidden barrier regions. We refer to this as resonant transport rather than resonant tunneling since no classically forbidden regions are involved.

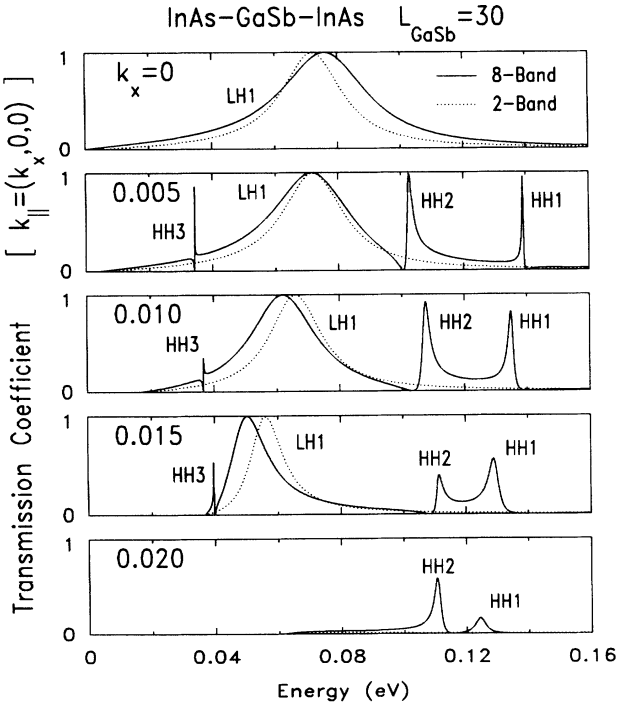

FIG. 4. Transmission coefficients for the InAs-GaSb-InAs structure calculated for a series of $\mathbf{k}_{||}$ values using the effective bond-orbital model (solid lines), and using a simple two-band model (dotted lines). The width of the GaSb quantum well is 30 monolayers, and k_x values labeling the plots are given in units of $2\pi/a$.

In Fig. 4 we show the transmission coefficients over the broken-gap energy range for an InAs-GaSb-InAs structure under flat-band conditions at several different values of $\mathbf{k}_{||}$. In addition to the transmission coefficients calculated using the eight-band effective bond orbital model (indicated by the solid lines), we have shown, for comparison, the results obtained with a simple two-band model[13] in which only the lowest conduction band and the light-hole band are included (dotted lines). At $\mathbf{k}_{||} = \mathbf{0}$, the two-band and the eight-band results are qualitatively similar, yielding a single light-hole transmission resonance peak which we designate as LH1. Since the formation of the quasibound state does not involve any barriers, the transmission resonance width is rather large—$\Delta E \approx 30$ meV, corresponding to an intrinsic quasibound state lifetime of ~ 20 fs.

For $\mathbf{k}_{||} \neq \mathbf{0}$, transmission coefficients calculated using the eight-band model exhibit a set of heavy-hole resonances not present in the two-band calculation. The heavy-hole resonances are considerably narrower than the light-hole resonance, indicating that the coupling between GaSb heavy-hole states and InAs conduction-band states is much weaker than that between the GaSb light-hole and InAs conduction-band states. The LH1 resonance is in general well approximated by the two-band model, although the eight-band model LH1 peak position decreases in energy more rapidly with increasing $\mathbf{k}_{||}$. The nonparabolic dispersion in the GaSb quantum-well band structure caused by heavy-hole–light-hole mixing is also evident in the position of the HH2 peaks as a function of k_x.

In Fig. 5 we show the probability densities (squared magnitude of the envelope function) near the GaSb quantum well for selected resonant transmitting states. The states shown have $\mathbf{k}_{||} = (0.01, 0, 0)$ and energies corresponding to the resonance peaks labeled HH1, HH2, and LH1 in the third panel of Fig. 4. For all three cases, the probability densities away from the GaSb quantum-well region are constant and consist predominantly of s components as expected for conduction-band plane-wave states in the InAs regions. In the GaSb quantum-well region, the probability densities consist mostly of the $J = \frac{3}{2}$ (heavy-hole and light-hole) components, and strongly reflect the nature of the quasibound states involved in the resonant transmission process.

The probability density of the HH1 state in the GaSb quantum well is dominated by the $J_z = \pm\frac{3}{2}$ components. The HH2 state shows large $J_z = \pm\frac{3}{2}$ components with a single node in the center of the GaSb well, which is a strong signature of the $n = 2$ heavy-hole state. However, the HH2 state also contains smaller but significant $J_z = \pm\frac{1}{2}$ components with no nodes, which is characteristic of

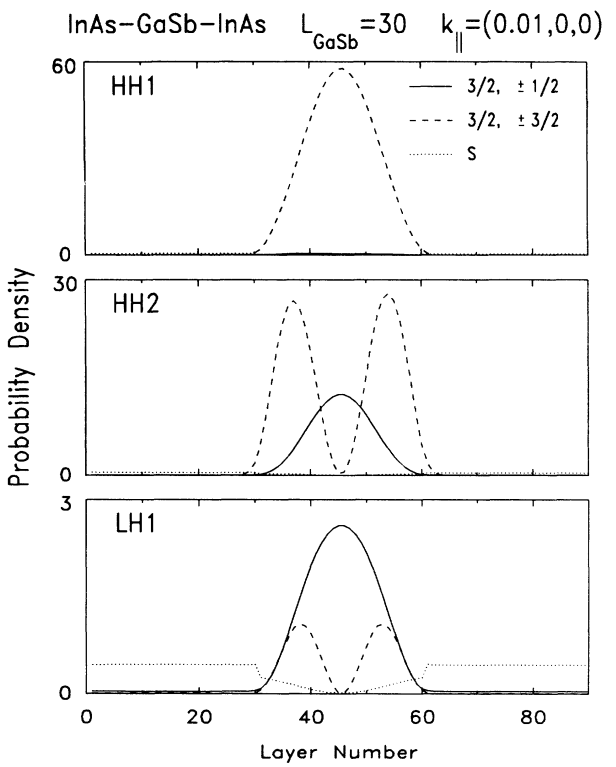

FIG. 5. Probability densities of selected transmitting states in the InAs-GaSb-InAs structure. The states shown have $\mathbf{k}_{||} = (0.01, 0, 0)$ and energies corresponding to the resonance peaks labeled HH1, HH2, and LH1 in Fig. 4. The GaSb quantum well is located at layers 31–60. The $(\frac{3}{2}, \frac{1}{2})$ and the $(\frac{3}{2}, -\frac{1}{2})$ components of the probability density are combined and shown as solid lines; the combined $(\frac{3}{2}, \pm\frac{3}{2})$ components are shown as dashed lines; and the combined s components are shown as dotted lines. The two $J = \frac{1}{2}$ components are insignificant and are not shown. Note that different vertical scales are used in the three panels of this figure.

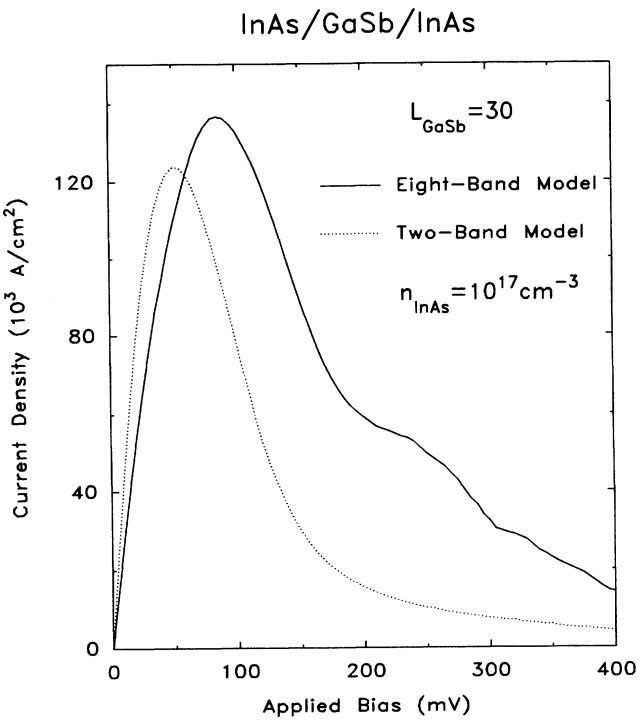

FIG. 6. Current-voltage characteristics for the InAs-GaSb-InAs structure calculated using the effective bond-orbital model (solid lines), and using a simple two-band model (dotted lines). The width of the GaSb quantum well is 30 monolayers. The doping level in the InAs electrodes is $n = 10^{17}$ cm^{-3}.

the $n = 1$ light-hole state. The LH1 state exhibits similar characteristics, except that $J_z = \pm\frac{1}{2}$ components are the largest. This indicates that both the states which we labeled HH2 and the LH1 contain mixed characteristics of the $n = 1$ light-hole and the $n = 2$ heavy-hole quasi-bound states. The mixing of the HH2 and the LH1 states can be readily understood in terms of the symmetries of the quasibound states. The $n = 2$ heavy-hole state has a p_z-like envelope function, while the $n = 1$ light-hole state contains p_z orbitals; both states therefore have overall p_z-like wave functions with respect to the center of the GaSb quantum well, resulting in the mixing of the states.

In Fig. 6 we show calculated current-voltage characteristics for the InAs-GaSb-InAs structure. Results from both the eight-band and the two-band calculations are shown. The eight-band result predicts somewhat higher peak current density and peak voltage, and exhibits a shoulder at 150 mV above the peak. We attribute this shoulder to the heavy-hole resonances. Our present calculation includes only elastic tunneling currents; in an actual device where inelastic-scattering processes also contribute to the current, the heavy-hole shoulder would be interpreted as a part of the valley current.

B. GaSb-InAs-GaSb

Figure 3(b) shows the GaSb-InAs-GaSb structure. This device is based on the same principle as the InAs-GaSb-InAs device, relying on the barrier-less resonant interband transport mechanism, with holes transported between the GaSb electrodes via the InAs conduction-band states. It may be regarded as the p-type version of the InAs-GaSb-InAs device. Figure 7 shows the $\mathbf{k}_{||} = (0.01, 0.01, 0)$ transmission coefficients in the broken gap energy range for a GaSb-InAs-GaSb structure under flat-band condition. Since both the incident and the transmitted waves can contain either light-hole or heavy-hole states, the transmission coefficient has four components. We denote the four components by LH-CB-LH, LH-CB-HH, HH-CB-LH, and HH-CB-HH. In all four components we observe the CB1 resonance corresponding to the lowest quasi-bound states in the InAs layer. The LH-CB-LH component provides the dominant contribution to the total transmission coefficient due to the strong coupling between the GaSb light-hole band and the InAs conduction band. The LH-CB-HH and the HH-CB-LH components each contribute equally to the transmission coefficient, totaling almost 20%. The HH-CB-HH component is quite small, although it shows the CB2 resonance not seen in the other components due to the fact that the CB2 resonance lies above the available GaSb light hole states at this $\mathbf{k}_{||}$ value.

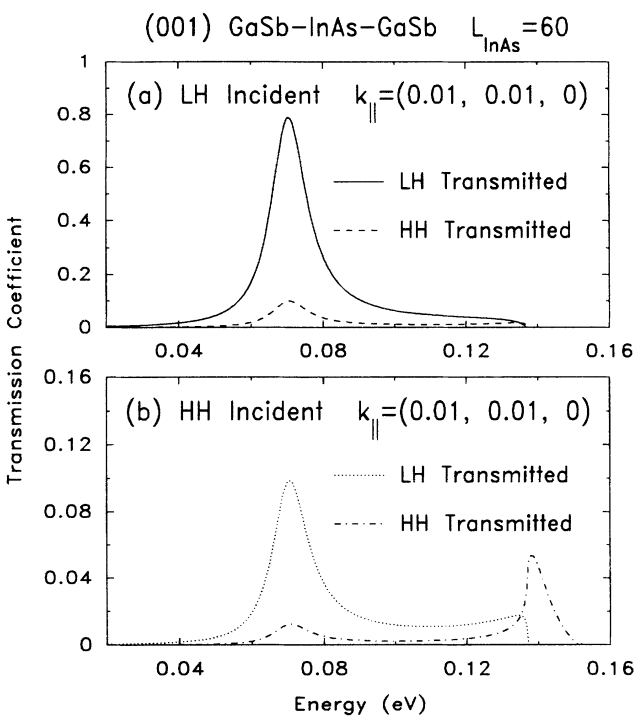

FIG. 7. Transmission coefficients calculated for the GaSb-InAs-GaSb structure for (a) an incident GaSb light-hole state and (b) an incident GaSb heavy-hole state. The width of the InAs quantum well is 60 monolayers, and $\mathbf{k}_{||} = (0.01, 0.01, 0)$. Note that different vertical scales are used in (a) and (b).

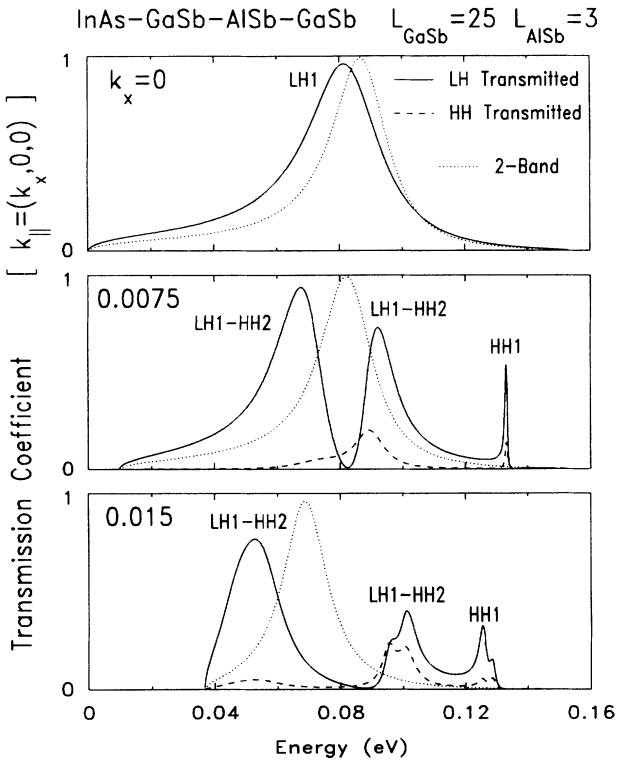

FIG. 8. InAs-GaSb-AlSb-GaSb transmission coefficients for a series of $\mathbf{k}_{||}$ values. The width of the GaSb quantum well is 25 monolayers, and the width of the AlSb layer is 3 monolayers. Both the two-band model results (dotted lines), and the eight-band model results (solid lines for LH transmitted, dashed lines for HH transmitted) are shown.

C. InAs-GaSb-AlSb-GaSb

In the final example we examine an interband device which has one n-type electrode and one p-type electrode. The most basic device of this type consists simply of an interface between n-type InAs and p-type GaSb,[8] which can be thought of as a heterojunction Esaki diode. Substantial improvement in the performance of this device can be obtained by inserting a thin AlSb barrier layer on the GaSb side of the interface, as depicted in Fig. 3(c), forming a GaSb quantum well and thereby inducing resonances for achieving enhanced tunneling.[10] In Fig. 8 we show the transmission coefficients for the InAs-GaSb-AlSb-GaSb device under flat-band condition for $\mathbf{k}_{||}$ values of (0,0,0), (0.0075, 0, 0) and (0.015, 0, 0), using both the two-band model (dotted lines) and the eight-band model (solid lines for transmitted LH states and dashed lines for transmitted HH states.) The widths of the GaSb quantum well and the AlSb barrier are 75 Å and 9 Å, respectively. At $\mathbf{k}_{||} = \mathbf{0}$, the light-hole transmitted component shows a single resonance peak identified as LH1, and there is no heavy-hole transmitted component. In this case the two-band model result approximates the eight-band result reasonably well. For $\mathbf{k}_{||} \neq \mathbf{0}$, a small but non-negligible heavy-hole transmitted component is found in addition to the light-hole transmitted component which still provides the dominant contribution. For $\mathbf{k}_{||} = (0.0075, 0, 0)$, two wide peaks and a single narrow peak are found. By studying the wave functions of the transmitting states, the narrow peak is identified as the HH1 resonance, while the two wide peaks are found to be resonances corresponding to LH1-HH2 mixed states. Because of the strong hole-mixing effects, the calculation for this structure using the two-band model compares very unfavorably to the more realistic eight-band model calculation. An additional feature in the InAs-GaSb-AlSb-GaSb structure is the lack of inversion symmetry. As a result, splitting of Kramers degeneracy is expected.[35] This can be seen at $\mathbf{k}_{||} = (0.015, 0, 0)$, where the larger dispersion makes the splitting clearly observable for the two higher peak structures. Note that no such splitting is seen for the symmetric InAs-GaSb-InAs structure at comparable $\mathbf{k}_{||}$ values (see Fig. 4).

Figure 9 shows the current-voltage characteristics for an InAs-GaSb-AlSb-GaSb structure calculated using the two- and the eight-band models. The peak current density in the eight-band J-V curve is more than three times higher than the two-band peak. At first it may be tempting to attribute this large discrepancy to additional currents due to heavy-hole resonances. A more careful analysis shows that this is false. In a diode with InAs and GaSb terminals, a transmission resonance contributes to the elastic tunneling current only if it lines up with both the InAs electron Fermi sea and GaSb hole Fermi sea in the two electrodes. As a result of this rather stringent requirement, the amount of elastic tunneling current in this type of device is very sensitive to the positions of

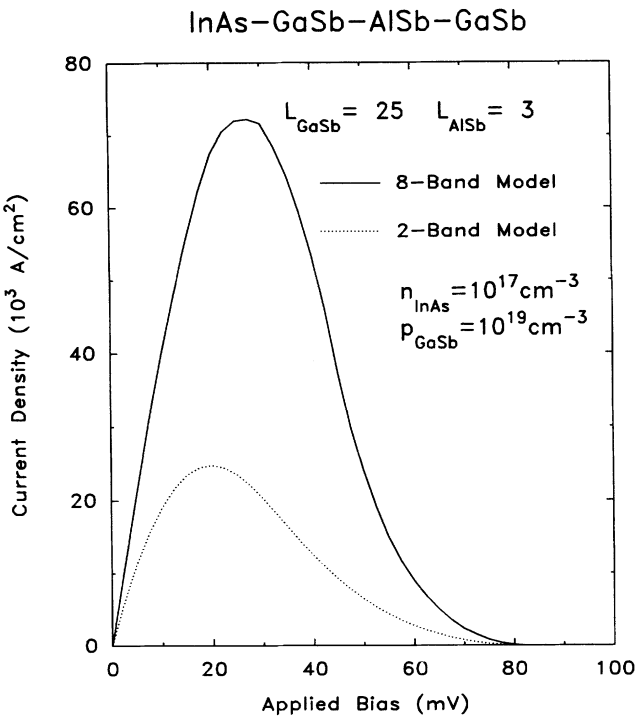

FIG. 9. Current-voltage characteristics for the InAs-GaSb-AlSb-GaSb structure calculated using the effective bond-orbital model (solid lines), and using a simple two-band model (dotted lines). The doping levels in the InAs and GaSb electrodes are $n = 10^{17}$ cm^{-3} and $p = 10^{19}$ cm^{-3}.

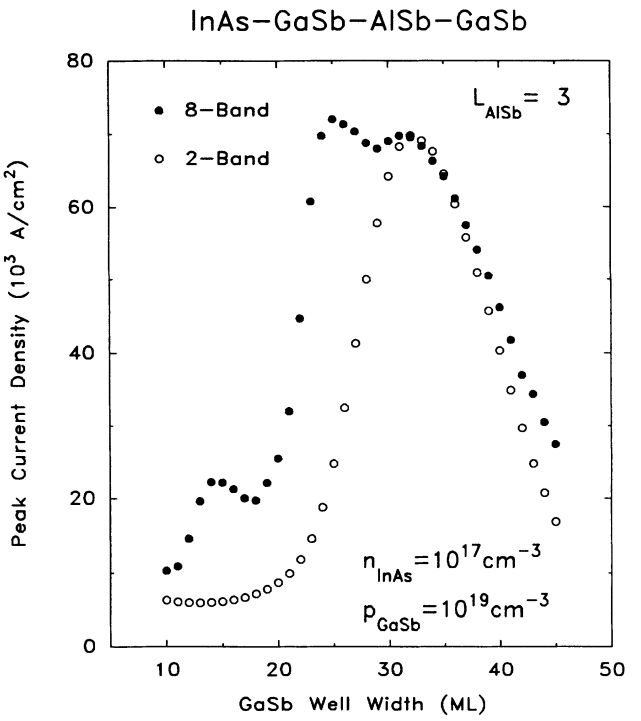

FIG. 10. Dependence of theoretical peak current densities on the GaSb well width for the InAs-GaSb-AlSb-GaSb tunnel structure, calculated using the effective bond-orbital model (solid circles), and using a two-band model (open circles). The width of the AlSb barrier is 3 monolayers.

the resonance peaks. As seen in Fig. 8, the nature of the LH1 resonance is drastically altered due to band mixing effects, splitting into two LH1-HH2 peaks. The difference in the hole transmission spectra accounts for the large discrepancy in the peaks current density computed with the two different models.

A broader view of the influence of heavy-hole states on J-V characteristics of InAs-GaSb-AlSb-GaSb tunnel structures is shown in Fig. 10, where we plot the two-band and the eight-band theoretical peak current densities as functions of GaSb well width. Due to the inclusion of heavy-hole states, the eight-band result contains considerably more structure. For narrow GaSb wells, the two-band model predicts no resonance in the broken-gap energy range, while the eight-band model shows the HH1 resonance. This is responsible for the higher peak current densities predicted by the eight-band model for narrow GaSb wells. For wider wells, the two-band model predicts a single wide light-hole resonance, while the eight-band model shows the split pair of LH1-HH2 wide resonances, as seen in Fig. 8. We see that while the overall magnitudes of the peak current densities predicted by the two models are similar, the eight-band model shows two local maxima rather than the one predicted by the two-band model.

IV. SUMMARY

We have studied InAs/GaSb/AlSb-based interband device structures using a recently developed method for computing transmission coefficients. Our method is a generalization of the method of Frensley[23] to multiband tight-binding models. With this method, the transmission probability can be calculated simply by solving a system of linear equations representing the tight-binding form of the Schrödinger equation over a finite region of interest, with specially formulated boundary and inhomogeneous terms to account for the effects of the incoming and outgoing plane-wave states.

Our method is numerically stable, efficient, and simple to implement, making it possible to compute transmission coefficients effectively in wide heterostructures using realistic multiband band-structure models. We have implemented this method for the effective bond-orbital model, which realistically describes the relevant band structure near the Brillouin-zone center. Comparing our calculations with results obtained using a simple two-band model which includes only the lowest conduction band and the light-hole band, we find that while the primary interband transport mechanism arises from the coupling between the InAs conduction-band states and GaSb light-hole states, in device structures containing GaSb quantum wells, the inclusion of heavy-hole states can introduce additional transmission resonances and substantial hole-mixing effects. We have calculated current-voltage characteristics for the InAs-GaSb-InAs and InAs-GaSb-AlSb-GaSb devices using both models, and have shown that the influence of the heavy-hole states could be substantial.

ACKNOWLEDGMENTS

The authors would like to thank W. R. Frensley, C. S. Lent, Y. C. Chang, and J. N. Schulman for helpful discussions. One of us (E.T.Y.) would like to acknowledge financial support from the AT&T Foundation. This work was supported by the Office of Naval Research (ONR) under Grant No. N00014-89-J-1141.

[1] G. J. Gaultieri, G. P. Schwartz, R. G. Nuzzo, R. J. Malik, and J. F. Walker, J. Appl. Phys. **61**, 5337 (1987).
[2] G. J. Gaultieri, G. P. Schwartz, R. G. Nuzzo, and W. A. Sunder, J. Appl. Phys. **49**, 1037 (1986).
[3] We assume transitivity to obtain the InAs/AlSb band offset value.
[4] J. R. Söderström, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **55**, 1094 (1989).
[5] L. F. Luo, R. Beresford, and W. I. Wang, Appl. Phys. Lett. **55**, 2023 (1989).
[6] R. Beresford, L. F. Luo, and W. I. Wang, Appl. Phys. Lett. **56**, 551 (1990); **56**, 952 (1990).
[7] K. Taira, I. Hase, and K. Kawai, Electron. Lett. **25**, 1708 (1989).
[8] D. A. Collins, E. T. Yu, Y. Rajakarunanayake, J. R. Söderström, D. H. Chow, D. Z.-Y. Ting, and T. C. McGill, Appl. Phys. Lett. **57**, 683 (1990).
[9] L. Yang, J. F. Chen, and A. Y. Cho, J. Appl. Phys. **68**, 2997 (1990).
[10] D. Z.-Y. Ting, D. A. Collins, E. T. Yu, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **57**, 1257 (1990).
[11] D. A. Collins, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **58**, 1673 (1991).
[12] J. R. Söderström, E. T. Yu, M. K. Jackson, Y. Rajakarunanayake, and T. C. McGill, J. Appl. Phys. **68**, 1372 (1990).
[13] D. Z.-Y. Ting, E. T. Yu, D. A. Collins, D. H. Chow, and T. C. McGill, J. Vac. Sci. Technol. B **8**, 810 (1990).
[14] E. T. Yu, D. A. Collins, D. Z.-Y. Ting, D. H. Chow, and T. C. McGill, Appl. Phys. Lett. **57**, 2675 (1990).
[15] D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Appl. Phys. Lett. **58**, 292 (1990).
[16] Y. C. Chang, Phys. Rev. B **37**, 8215 (1988).
[17] E. O. Kane, in *Semiconductors and Semimetals*, edited by R. K. Willardson and A. C. Beer (Academic, New York, 1966), Vol. 1, p. 75.
[18] E. O. Kane, in *Tunneling Phenomena in Solids*, edited by E. Burstein and S. Lundqvist (Plenum, New York, 1969), p. 1.
[19] J. N. Schulman and Y. C. Chang, Phys. Rev. B **27**, 2346 (1983).
[20] C. Mailhiot and D. L. Smith, Phys. Rev. B **33**, 8360 (1986).
[21] G. Wachutka, Phys. Rev. B **34**, 8512 (1986).
[22] D. Y. K. Ko and J. C. Inkson, Phys. Rev. B **38**, 9945 (1988).
[23] W. R. Frensley (private communication).
[24] T.B. Boykin, J.P.A. van der Wagt, and J.S. Harris, Jr., Phys. Rev. B **43**, 4777 (1991).
[25] J. N. Schulman (private communication); J. N. Schulman and D. Z.-Y. Ting, Phys. Rev. B (to be published).
[26] C. S. Lent and D. J. Kirkner, J. Appl. Phys. **67**, 6353 (1990).
[27] D. Z.-Y. Ting, E. T. Yu, and T. C. McGill, Phys. Rev. B **45**, 3576 (1992).
[28] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).
[29] M. K. Jackson, D. Z.-Y. Ting, D. H. Chow, D. A. Collins, J. R. Söderström, and T. C. McGill, Phys. Rev. B **43**, 4856 (1991).
[30] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986), pp. 578–614.
[31] C. B. Duke, *Solid State Physics*, Tunneling in Solids, Suppl. 10 (Academic, New York, 1969).
[32] Y. C. Chang, Phys. Rev. B **25**, 605 (1982); Y. C. Chang and J. N. Schulman, *ibid.* **25**, 3975 (1982).
[33] J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart, *LINPACK User's Guide* (SIAM, Philadelphia, 1979).
[34] See, for example, G. Dahlquist and Å. Björck, *Numerical Methods*, translated by N. Anderson (Prentice-Hall, Englewood Cliffs, NJ, 1974).
[35] See, for example, P. T. Landsberg, in *Solid State Theory; Methods and Applications*, edited by P. T. Landsberg (Wiley, London, 1969), p. 506.